

## 地球化学におけるデータ駆動科学

鳥海光弘\*

(2013年4月1日受付, 2013年5月16日受理)

### Data-Driven Geochemistry in Solid Earth Processes

Mitsuhiro TORIUMI\*

\* Japan Agency for Marine-Earth Science and Technology (JAMSTEC)  
2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan

Recent investigations of geochemical processes in the earth's interior produce continuously huge amounts of various data. Many geological and geophysical processes can be investigated by recently developing data-driven machine learning methods to make realistic modeling combining geochemical, geological, geophysical, and seismological time series data. In this paper, the author tries to apply the dimension reduction using phase-reduction equation method for modeling the geochemical processes in the plate dynamics.

**Key words:** Plate dynamics, Geochemistry, Big data, Dimension reduction, Complexity

### 1. はじめに

さる2011年3月11日に東北日本から関東地方を襲った巨大境界型地震と、それ以前の100年間に発生した太平洋およびインド洋における超巨大地震群により、人間社会は大きな被害に遭遇した。そしてさらに、巨大地震が、さらなる巨大火山噴火などの自然災害とどのようにリンクしているかについて、地球化学のもつ長期にわたる時間変化の測定可能性を有効利用する手法の開発が急がれてきた。

プレート力学では、プレート境界領域の地下構造研究や、特徴的な断層運動にもなう流体分布やガス放出などの地球化学的な研究など、そして島弧マグマやプレート境界の変成岩の成因に関する岩石学、あるいは地球化学的研究が中心的なものであった。しかし、実践的な化学としての地球化学が、火山噴火や巨大地震という短時間現象にも、物質の変化や変化の過程を記録し、読み取る科学として他の追従を許さない領域であることから、その社会への責務が大きいものであ

るということがより強く意識されるようになった。ましてや、東アジア全域でひろまってしまった原子力発電所とその使用済み燃料棒などの安定的な隔離と保全には、地球化学が重要な役割を担わざるを得ない。

さらに重要な契機があった。それは2010年に始まったビッグデータ利用に関する科学と工学分野全域にまたがる大きなうねりである。はじめはヒトゲノムと遺伝子表現との膨大な関係性の解読である。そしてさらに拍車をかけたのは、インターネット上の巨大で多様な情報をサーチし、分類し、関連づけをおこなうことや、それを介してスマートフォンによって個々人の24時間の行動が記録され、それがたとえば関東圏では3000万人分という膨大な量になっていることに端を発していた。

つまり、そのような場面では10の8乗以上のデータ次元が時間軸に有限個の時系列を作るのである。それがたとえば10の5から7乗の時系列データであるなら10の15乗という量になる。それを解析する上では相関係数をとるのであるから、もはや計算不能問題となる。それがビッグデータまたはデータ洪水の本質的な問題なのであった。

\* 独立法人海洋研究開発機構地球内部ダイナミクス領域  
〒237-0061 横須賀市夏島町2-15

## 2. 地球化学における縮約位相方程式

ところで、地球化学におけるビッグデータはどのようなものであろうか。地球表層から内部を巻き込んだ、多数の物理化学過程と生物過程の結合による主成分元素、微量成分、極微量成分、有機物、イオン種、ガス成分、同位体成分などがそれぞれ高次元ベクトルデータをつくっている。そしてその高次元ベクトルデータ空間は、様々な物理化学的あるいは生物的過程が実在空間に配置され、その結果、空間次元と地球化学ベクトル空間からなる情報空間上で、時系列データが、物理化学過程と生物学過程（併せて、地球化学過程とよぶ）に支配された部分空間を構成しているのである。つまり、次のように位相方程式を考えるのである。

はじめに次の連立1次非線形常微分方程式を考える。

$$dX/dt = f(X)$$

$$X = X(X_1, X_2, X_3, \dots, X_n)$$

地球化学過程では化学成分が  $X/abs(X)$  にあたり、これが空間的に、また時間的に非周期的な時間変動をしている。ただし、 $abs(X)$  は  $X$  の絶対値である。それぞれのベクトル  $X$  に対して、スカラー  $g = g(X)$  を決

める。このとき  $g$  は

$$dg/dt = w + noise$$

に従うとする。このとき  $w$  は

$$w = (grad\ g \cdot f(X))$$

ただし、 $grad$  は  $X$  に関する  $g$  の勾配であり、 $( )$  はベクトル内積である。

$X$  は高次元空間上で、非周期的な変動をしているので、スカラー  $g$  は  $X$  変動の位相となる。つまり  $w$  は位相方程式の角速度に対応したものである。Fig. 1では、 $X$  座標上の超平面にほぼそって変動する点群が観測時系列のつくる軌跡であり、それは超平面上のアトラクタとなっている。これは、アトラクタ中心付近の固定点から  $X$  へのベクトルの角度を  $g$  としてその時間変化率を  $w$  としたことに対応している。むしろ、線形の周期系ではないので、 $g$  や  $w$  は  $X$  によって変化するものであろう。このような位相ダイナミクスの導入は蔵本 (2005) に詳しい。

こうした超平面上で、時空間に配置された点の軌跡である時系列データは、それらの集団的軌跡が超平面部分空間のなかで特有な軌道、すなわちアトラクタをつくり、それこそが地球化学過程のダイナミクス方程式、ここでは低次元に縮退した非線形常微分方程式となり、アトラクタ上の縮約位相方程式となる。こうし

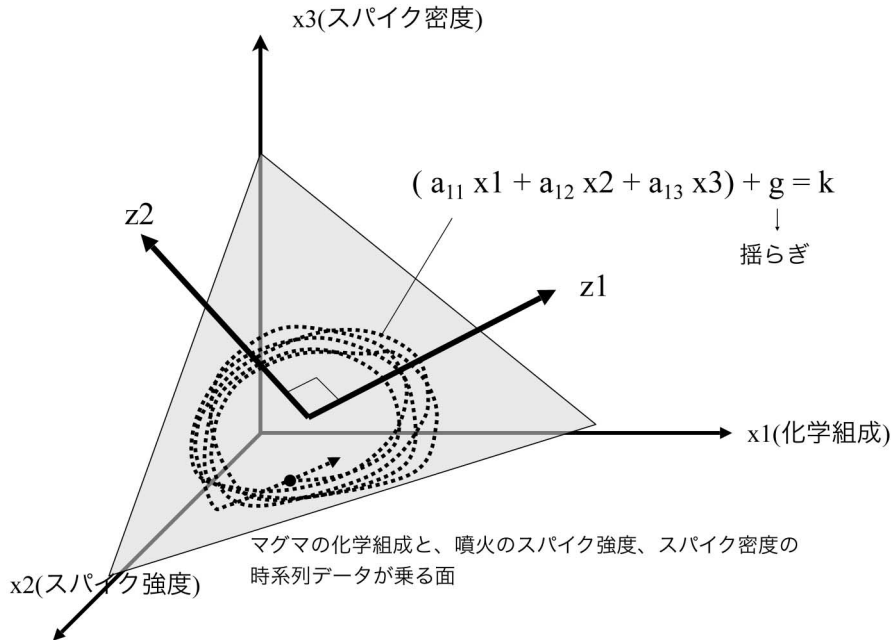


Fig. 1 Concept image of the geochemical and geophysical data space and reduced super-surface by the dimension reduction method. Time series of data related volcanic activity are shown as the attractor on the super-surface.

て得られる位相方程式はむしろ揺らぎがあるので、確率微分方程式で表現される。そして、縮約された位相方程式は、一つの位相ではなく、複数の位相をもってよい。その場合、アトラクタは高次元トラスを形作る。

### 3. 地球化学におけるビッグデータの次元圧縮

地球化学上の同位体や各種元素成分量のベクトル空間では、基底はそれぞれの規格化した組成であるが、物理化学過程に応じて対数関数や指数関数で変換されて、互いに従属関係にある。たとえば、最大分別作用でマグマからかんらん石や斜長石を結晶化させてマグマの化学組成を変化させると、いろいろな成分で分配法則が成立して、時間を消去できる。そのときの多成分の間の関係がここでいうノイズを含む位相方程式である。すると、ある物理的過程の起こった空間的な領域のなかで、多数の観測点での、地球化学的データおよび物理的変動のデータの時系列データ群は、その空間的領域における物理化学過程を支配しているダイナミクスに対応した、超平面上のアトラクタを持つはずである。ただし、それは取り上げられた地球化学変数のベクトル空間ではなく、そのなかの縮約された部分空間上のアトラクタである。こうして得られたアトラクタが、大量のビッグデータからの主成分解析により次元圧縮により抽出された新しい変数空間のアトラクタであることは容易にわかる (Bishop, 2006参照)。こういう方法で法則性を抽出するのはデータ駆動型科学と呼んで、分野を超えて拡がった一般的科学である。

さて、以上のようにして求められたアトラクタは、分配法則や最大分別などから導かれる関係の時間変動とは違ったものとなる。それはいくつかの分別過程や力学過程が競合して起こり、必ずしも一定の割合で進行しはしないからである。つまり、時間的にも空間的にも不均一な過程だからである。そして、重要なことは、物理化学過程が化学的データだけでない。たとえば、火山活動のダイナミクスについて考えてみよう。火山をつくる地下のマグマ溜まりは、地殻やマンツルの応力と歪場を変化させる。とくに火山フロントは、地下のマグマ溜まり、さらに深部のマンツルの部分融解層では、マグマは様々な程度で連結しているために、その連結様式によって異なる周期的な変動の遅れ、つまり位相の遅れが現れるだろう。個々の火山と多数の火山のいろいろな観測変数の時系列データ、お

よび火山フロントの地殻変動や微小地震発生率などの時系列データから、それらのダイナミクスの縮約位相方程式を求めることによって、マグマ活動と、火山フロントからプレート境界領域の地殻とマンツルの地殻変動との力学系が再構築できるだろう。

### 4. 火山噴火のダイナミクスと位相縮約

火山活動にともなう、マグマの温度、体積、化学成分の変動などに加えて、地殻変動や、磁場変動、さらには微小地震活動なども同時に取り扱うことができる。このような異質なデータを単純な線形結合でつなげるのは無理がある。そのため、化学組成は対数関数に変換し、傾斜計のデータは体積歪みに変換、他のデータも非線形変換したうえで、それらの線形結合を取り扱うことにする。Fig. 1に示したように、ある活動を始めた火山領域について、多数の物理化学データを変数変換した上で、それらの変数の高次元ベクトル空間上の、縮約した部分空間を、主成分解析による、いくつかの主成分空間から求める。Fig. 1には、元の高次元空間上の多数のデータが乗る超平面が示されている。この次元圧縮された超平面上に多数の時系列データを表現すると、図のような軌跡となり、アトラクタをえる。この超平面上のアトラクタを重心点から各点を結ぶベクトルの角度変数を位相角とすると、この火山活動の縮約位相方程式を得たことになる。

このような解析を行うことは、システムの挙動を統合的に明らかにすることになる。ひとつの火山では地下にあるマグマ溜まりのサイズとそこへのマグマの流入と噴火による流出、それらのマグマの化学組成、マグマ溜まりでの結晶分化とその結果によるマグマの層状構造と対流構造、そして、温度構造とガス圧力が重要なシステム変数となる。それらは時間変動を示す。それらの物理過程と化学過程のさまざまな変数を同時に取り扱い、十分な数の時系列を測定することで、火山噴火のダイナミクスについてデータ駆動型の物理化学法則を捉えることができるだろう。

ここで問題となるのは、化学変数や物理変数が時間と空間に対して均質なデータを取得できないことである。火山噴火が間欠的であり、時間間隔も一定ではない。そのような非周期のデータ取得に対応した補完をおこなうことが必要である。噴火が地震と同じようにスパイクな事象であり、スパイク噴火で放出されるマグマやガスの化学組成が、連続的な隠れたプロセスの産物であっても、噴火という物理過程に従属している

に違いなく、したがってスパイク強度やスパイク密度も重要なシステムデータであることである。

Fig. 1には、化学変数とスパイク密度つまり噴火の時間間隔の逆数、及びスパイク強度つまり一回の噴火のマグマの量を、元の物理化学系の基底として、主成分解析により次元圧縮して作られた超平面上で時系列がとる軌跡、アトラクタが点線で表現されている。実際の火山噴火現象がこのように解析できるかは明らかではない。

本文は仮想的な研究報告を自由形式で書いた。実際の研究についてはまだないが、すでに構想段階は過ぎて実際のデータについて検討中であり、2030年というより、2015年が相応しいかもしれないこと附記したい。

#### 引用文献

- Bishop, C. M. (2006) Pattern Recognition and Machine Learning. Springer-Verlag, NY, pp. 760.  
蔵本由紀 (2005) リズム現象の世界. 東京大学出版会, pp. 242.